TO STUDY THE DRUG TARGET IN EBOLA VIRUS AND EXPLORE ANTI-VIRAL AGENT *IN SILICO*

A thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Bachelor of Technology

in

Biotechnology

Submitted by

Sonit Kumar Jena

Under the Guidance of

Dr. B.P.Nayak

Department of Biotechnology and Medical Engineering

National Institute of Technology, Rourkela, India

2015

Department of Biotechnology and Medical Engineering

National Institute of Technology

Rourkela

CERTIFICATE

This is to certify that the thesis entitles **"To study the drug target in Ebola virus and explore anti-viral agent *in silico* "** submitted by **Mr. Sonit Kumar Jena** in partial fulfilment of the requirements for the degree of Bachelor of Technology in Biotechnology in the department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela is an authentic research work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in this report has not been submitted earlier to any university/ Institute for the award of any Degree or Diploma.

Date :

Place : Rourkela

Dr. Bibhukalyan PrasadNayak

Assisstant Professor

Department of Biotechnology and Medical Engineering

National Institute of technology, Rourkela-769008

CONTENTS

Page No.

ACKNOWLEDGEMENT

I take this opportunity to express my hearty gratitude and profound regards to my thesis supervisor, **Dr.B.P.Nayak**, Assistant Professor, Department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela, for whose constant guidance and encouragement led to the successful completion of my B.Tech thesis. It is for his esteem guidance and noble supervision which resulted in the materialization of this project. I would also thank him for helping me to infer my mistakes objectively and inspiring me to improve upon it , thus inculcating a scientific temperament and keeping my interest alive in the subject as well as for being approachable at all the times.

I would also extend my sincere thanks to Mr. Binay Kumar and Mr. Rakesh Bhulan, M.Tech student for providing me knowledgeable insight regarding Bioinformatics tools which were of much relevance for my project work.

I am also thankful to all the lab-members working under the supervision of Dr.B.P.Nayak for their suggestions and ideas which resulted in several modifications in my data and helped me set up a better presentation.

Finally, I am also grateful to all my batch-mates for keeping up an intellectual environment in the department and their support and unity which kept us patient during difficult times.

Sonit Kumar Jena

Department of Biotechnology and Medical Engineering

List of Abbreviations

1. EBOV : Ebola Virus
2. EVD : Ebola Virus Disease
3. PDB : Protein Data Bank
4. TIM-1 : T Cell immunoglobulin mucin domain-1
5. RNA: Ribonucleic Acid
6. VP: Viral Protein
7. GP : Glycoprotein
8. ADT : AutoDock Tools
9. ACD : advanced chemistry development
10. DS : Discovery Studio
11. ADMET:Absorption-Distribution-Metabolism-Excretion-Toxicity

List of Figures

List of Tables

ABSTRACT

Ebola virus is a virulent pathogen that causes a highly lethal hemorrhagic fever in human and nonhuman species. It has a lethality rate of around 50-90%. Ebola Virus is a single stranded & negative RNA virus. It is a pathogen that belongs to Filoviridae family of RNA virus. Ebola is considered as a level-4 agent as it requires high range of biohazard containments for analysis and studies. The rapid growth of this virus infection has made the scenario so complicated to control this disease. Ebola virus is unreceptive to many anti-viral drugs or antibiotics also not a valid treatment is yet discovered for this disease caused by this virus. So there is an urgent need to explore the drug targets and discover novel anti-viral agent to fight against the virus. Proteins that used to help in pathogenesis of organism are selected as drug targets. Subsequent researches identified a cell surface receptor called TIM-1 for Ebola and Marburg viruses that typically enhances filovirus infection in many cells. Various function of TIM-1 in pathogenesis made it a potential drug target to control infection. Using bioinformatics based approach potential drug molecules are discovered against the target protein. The objective of the current study to explore a novel anit-viral agent that does act against TIM-1 receptor of Ebola virus using bioinformatics tools. Briefly, 25 inhibitors against TIM-1 protein were selected from literature. Both ligand and protein were optimized using Autodock Vina software. TIM-1 was docked with those selected compounds and stability was estimated on the basis of total binding energy. Deslanoside, Vinorelbine, Digoxin were found to be most suitable agent that inhibit TIM-1. Further, drug likeness and ADMET properties were evaluated by using molinspiration & Pre-ADMET. These compounda can be taken for wet lab testing as anti-Ebola therapy. *In-silico* docking analysis is done repeatedly to find the most suitable compound.

Keywords: *In silico,* hemorrhagic fever, TIM-1, EBOV, Deslanoside , Venorelbine

CHAPTER 1

INTRODUCTION

1.INTRODUCTION

EBOV is a deadly pathogen belonging to Filoviridae family of RNA virus which is single-stranded and negative RNA virus. It typically cause hemorrhagic fever in human and non-human species with mortality rate very high (50-90%). Hemorrhagic fever mainly leads to death and severe bleeding from different parts of the body [1]. It takes place in two phases, late phase & incubation period. The natural host of EBOV is still not known. The rapid growth of this virus infection has made the scenario so complicated to control EVD. The specific mechanism of the pathogenicity of EBOV has not been clearly discovered yet. No current anti-viral drug or anti-viral therapy has been invented which could be effective against the infection. A high range of biohazard containments are required for research studies and analysis. EBOV is termed as a level 4 agent which alludes that it is extremely infective and lethal.

EBOV was first seen in Africa in 1976 in the republic of Congo. Outbreaks have been appeared for past 3-4 years in central Africa. Outbreak in 2014 is one of the largest outbreaks in some places of West Africa [17]. Countries like Guinea and Sierra Leone have been severely attacked with the syndrome of Ebola Virus according to the report Centre for Disease Prevention and Control. But still there is no conventional therapeutics has been discovered to inhibit the infection of this deadly virus. Ebola infection has been increasing rapidly in an exponential way around the globe. From the statistical analysis it has been found that the lethality rate has reached to the extreme figure (83%). Though the virus source is still not known, the mode of transmission is well described [2]. The therapy to combat the infection

due to this virus has been a challenge. Drugs are very less available and limited. So some screening approach to find approved drugs which can be proposed as a potential anti-viral agent for the treatment of patients with EBOV infection. So mainly Ebola virus cause hemorrhagic fever as the virus separates the walls of the capillaries, which leads to leaking of blood from different parts of body. Basically hemorrhagic fever occurs due to 5 RNA virus family including Filoviridae family. So there is an urgent need to develop therapeutic domains. *In silico* methods have been adopted for screening of different compounds and to develop an anti-viral agent to fight against several disease. Computer added drug design is one of the best bioinformatics approach to develop selected drug leads which can be proposed further for clinical analysis of different disease. These computer and computational methods involve protein structure to be docked with individual drug molecule. Further calculating the binding affinity scores, the best drug lead can be proposed. This study deals with docking of 25 inhibitor molecules for Ebola. Consequently a series of docking has been performed with the ligand and the drug target. Based on the dock score or binding energy values with some other properties, the potential drug leads have been obtained. The main aim of this study to contribute some effort for researches, those are being carried out in clinical section. So this hypothetical approach can be helpful for molecular therapeutic analysis and studies to get a detailed effect of therapy for healing the infection.[2] [4]

CHAPTER 2

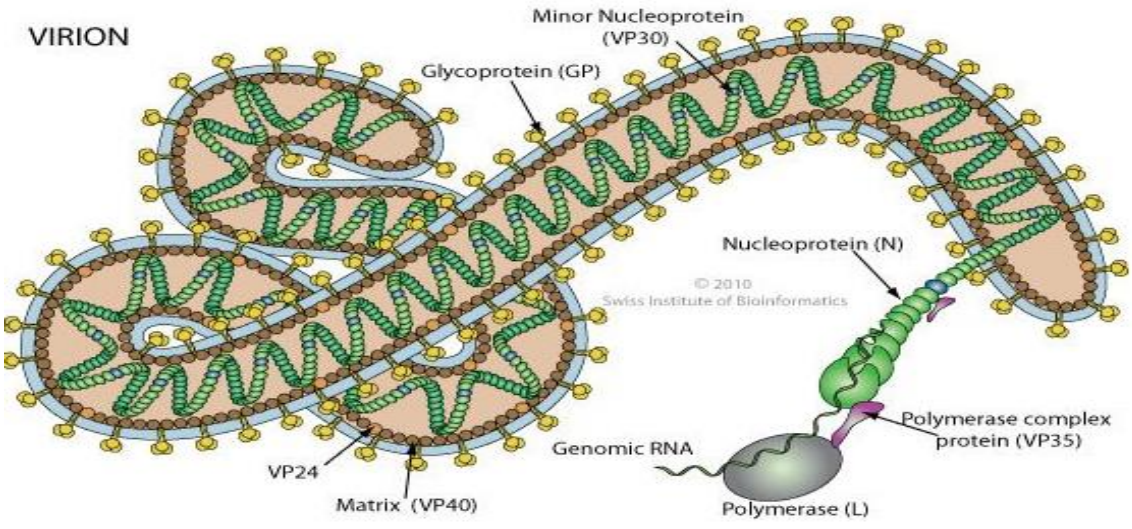
LITERATURE REVIEW

2.LITERATURE REVIEW

2.1. EBOLA VIRUS:

2.1.1. Structure:

Ebola Virus is a single stranded (ss) negative RNA virus which is in encapsulated form. This virus comes under the family of Filoviridae. It is having a highly filamentous structure. It looks like it has a thread like structure. It may be coiled or branched. EBOV is typically 800nm long and diameter of around 80nm. Each virion contains a nucleocapsid with in which a negative ss RNA is present. The genetic material RNA is covered by nucleoprotein (NP). Ebola has also a enzyme called Polymerase that helps for replication. Hence the RNA is surrounded by NP, Polymerase and four viral proteins (VP40, VP35, VP24, and VP30). The nucleocapsid is encapsulated by outer viral envelop. In between the matrix of nucleocapsid and viral envelop VP40 and VP24 are present.[4] [5]



(Figure 2.1virion structure showing all its components)

EBOV genome is around 10kb length with seven ORF that encodes for all structural proteins like VP35, NP, VP30 and Polymerase. It also encodes for matrix associated protein (GP, VP40 and VP24). Like the Marburg virus, the Glycoprotein of EBOV gives two gene products, one is a protein of 60-70 kDa and other one is of 150-170 kDa. These get into membrane by Transcriptional editing.[4] [15]

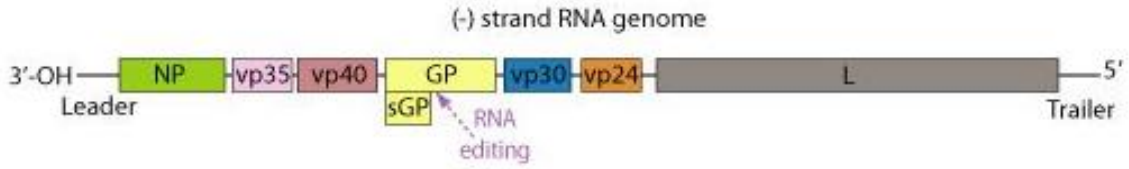


Figure 2.2 Gene order of the virus genome

Matrix Proteins:

EBOV viral protein (matrix protein) VP40 does the shaping and regulates the process for the purpose of budding. In VP40, subunits are merged to get a hexamer structure for this viral protein. A linker which is flexible is in contact at the ending part of chain. Structural integrity has to be maintained inside the virion. There is an association of budding of virus and endocytosis. Another matrix protein VP24 has a role to lower down the production of interferon. The important thing is the correct combination of NP, VP35 and nucleocapsid. [3]

Nucleocapsid:

At the mid portion of the virus, the nucleocapsid is present. It has a role of protecting the genome. The RNA is wrapped by the Nucleoprotein (NP). It creates a complex structure which is helical mostly, but it is not rigid. So EBOV looks like a wave structure. As the EBOV enters in to the cell, the polymerase creates much more number of copies of genome.[18] [19]

Glycoprotein:

Glycoprotein (GP) does the RNA editing and on cell surface protein is expressed like GP Spikes. These help to fuse the membrane of the viral particle into the cell. This is important for EBOV pathogenicity. [16]

2.1.2. Ebola Virus Transmission:

By direct interaction with bodily fluids such as blood, sweat, vomit, urine etc., it really cause to transmit the virus into human. It is one of the primary reason for outbreak of EBOV. Fruit Bats seemed to have this syndrome first. EBOV infection first appeared in different such as bat, monkey etc. To replicate in the animal cell with causing a high infection is the main aim or goal of Ebola virus. Ebola Virus contain GP and now the cell recognize the GP. The glycoprotein receptor on the cell surface get attached with the virus containing GP. This is the way Ebola virus enter the cell. Once Ebola virus enters the cell, it will release its contents, the genetic material, nucleoprotein & polymerase. Thus the genetic material will be replicated and undergoes some few processes like transcription and translation, which will create other structure that that the virus requires. The virus basically hijack the cell and replicates within it. Once it replicates it's all contents (genetic material, polymerase (L), nucleoprotein (NP)) will be packaged up again to create multiple Ebola virus. So the virus keeps replicating inside the bat cell. The bat cell is the optimal environment for Ebola virus replication. Now the infected fruit bat can transmit the Ebola virus to a human. Infection is by direct contact from blood or other bodily fluids.[4] [6] [8]

Incubation period:

It is the time required from the infection with the virus to which a symptom occur which is in between 2-21 days. Different methods adopted for diagnosis of EBOV such as ELISA, antigen recognition analysis, PCR.

2.2. DRUG TARGET FOR EBOV:

2.2.1. TIM-1:

Proteins that used to help in pathogenesis of organism are selected as drug targets. Subsequent researches identified a cell surface receptor called TIM-1 for Ebola and Marburg viruses that typically enhances Filovirus infection in many cells. Various function of TIM-1 in pathogenesis made it a potential drug target to control infection. It is one cell surface protein or receptor which essentially bind to the GP of EBOV. The Igv domain of this receptor contain pdtser are crucial for EBOV entry. As the pdtser interacts with the envelope of virion, it helps to entry of the virus. Hence it causes the infection inside. Pdtser is basically a residue for binding purpose or in other way it is a binding residue of Igv. TIM-1 increases basically infection in many cells such as epithelial cells. Ebola virus GP direct interacts with it and spread infection. [9]

Recently found TIM-1 (also having some other name: HAVCR-1, KIM-1) mainly expressed on epithelial cells, B cell, mast cell etc. It behaves as a receptor on the surface. The immunoglobulin (Igv) variable which is one type of domain, they are actually extended from PM. There is one another cell surface receptor called as niemann-pick (NPC1) which also seem to be required for EBOV entry into host cell. So both NPC1 and TIM-1 can be the potential drug target for an anti-ebola drug. But these NPC1 are not expressed in all type of cells. So mainly the Ebola infection is done by the other one i.e. TIM-1. Designing of one anti-viral drug that could

be able to restrict the virus entry into cells like epithelial cell. This anti-viral molecule somehow reduce the effect of the disease. TIM-1 essentially binds to virus GP and allows EBOV to enter to get infection in permissive cells. Researchers found this receptor which is the first cellular receptor for EBOV. It is the receptor which is identified first. So in this study, the drug target has been taken as TIM-1, which is to be docked with different drug molecules to design our anti-viral agent of interest. [11]

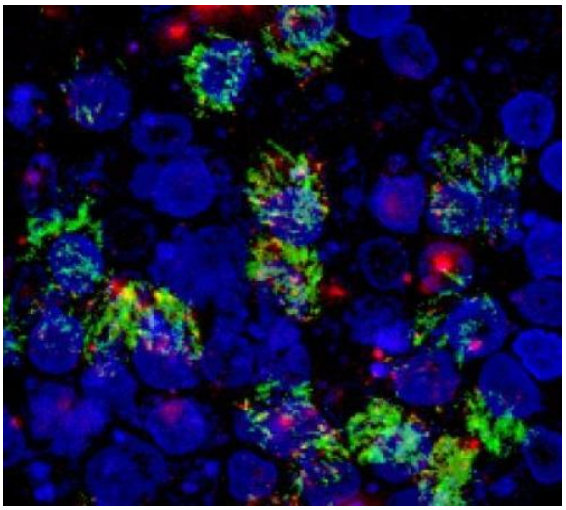


Figure 2.3 TIM-1 (in green) expressed on human cell, microscopic image

CHAPTER 3

TOOLS REQUIRED FOR IN SILICO ANALYSIS

3.TOOLS REQUIRED FOR *IN SILICO* ANALYSIS

3.1. PDB:

PDB (Protein Data Bank) is one of the single and best database which keeps the record of all data of proteins, genes, biomolecules, and nucleic acids. It was first introduced in Brookhaven National Lab. in the year 1971. It keeps all the structural data and info. of biomolecule which are deduced from XRC as well as NMR technology. It can be freely used and can be accessed on the internet referring to the site www.pdb.org.

3.2. PUBCHEM:

This is the database which carries all chemical molecules and it also contains its activities opposing to the biological analysis. This database is regulated by NCBI which is a part of NLM. Consequently this has been maintained by NIH. It can be easily accessed in the internet and the structures of all macromolecule can be freely downloaded from the website. Other properties of molecules can be checked like MW, H-bond donor and acceptor and logP values etc.

3.3. AUTODOCK TOOLS:

Autodock is a free software which can be used for the purpose of docking study. It can predict and measure the molecule like drug compounds which bind to the protein or receptor with 3D structure. So it is an automated tool for predicting the ligand and receptor interaction with showing different free affinity values. Different versions of these software are

available online, which can be freely downloaded from the site (www.autodock.scripps.edu). In the field of CADD for designing required drug compounds, it plays a vital role to help in developing many bioactive compounds. There are a lot of progress going on in XRC technology which allows main protein structures as well as nucleic acids also. These can be acted as agents for the control of different diseases that occurs in primates. The study of binding of such agents with potential target is crucial for development methods.

3.4. CHEMSKETCH:

Chemsketch is a software block from ACD, from which it can help biochemist and researchers to draw the chemical structures of different compounds, drawing schematic chemical diagrams, calculation of properties such as chemical property and developing analysis reports etc in a quick manner.

ACD/ChemSketch follows some steps:

- Mode is Structure so as to draw the structure in chemical form and to calculate different properties of molecule.
- Mode is draw or text and processing in graphic way.
- Modules that are additional can extend the possibilities of ACD/chemsketch.

3.5. CASTP SERVER:

CastP server is a bioinfo. tool used to carry out the prediction of active site of required protein. It also predicts geometric location of protein as well as delineation and measurement of surface areas of 3D structure of corresponding protein. Protein structures downloaded from PDB are also used here to find the pocket information, area of pocket, circumference and surface area of openings of the pocket. This server is regulated by Illinois University, Chicago. [12]

3.6. OPENBABEL:

Openbabel is a free software which is used for the purpose of changing file format from one to another. It is a very easy and fast exchangeable tool as it takes very few time for conversion of the file format. In autodock tools it demands the file format of both protein and ligand in .pdb format, so mainly openbabel converts any format to .pdf format in a quick process.

3.7. PRE-ADMET:

PreADMET is an application which is web-based and can be used for prediction of ADME readings. In silico method usually used to develop drug likeness record. PreADMET has a current version 2.0 which has four publications and it is available commercially. From PreADMET different properties can be calculated like drug likeness, ADME and analysis of toxicity.

3.8. PYMOL:

Pymol is a visualization software created by WL Delano and it is an open source tool. It creates high quality visualization of biomolecules with 3D structure. It is recently commercialized by Schrodinger.inc.It is vastly used in the field of computer based drug designing and structural biology.

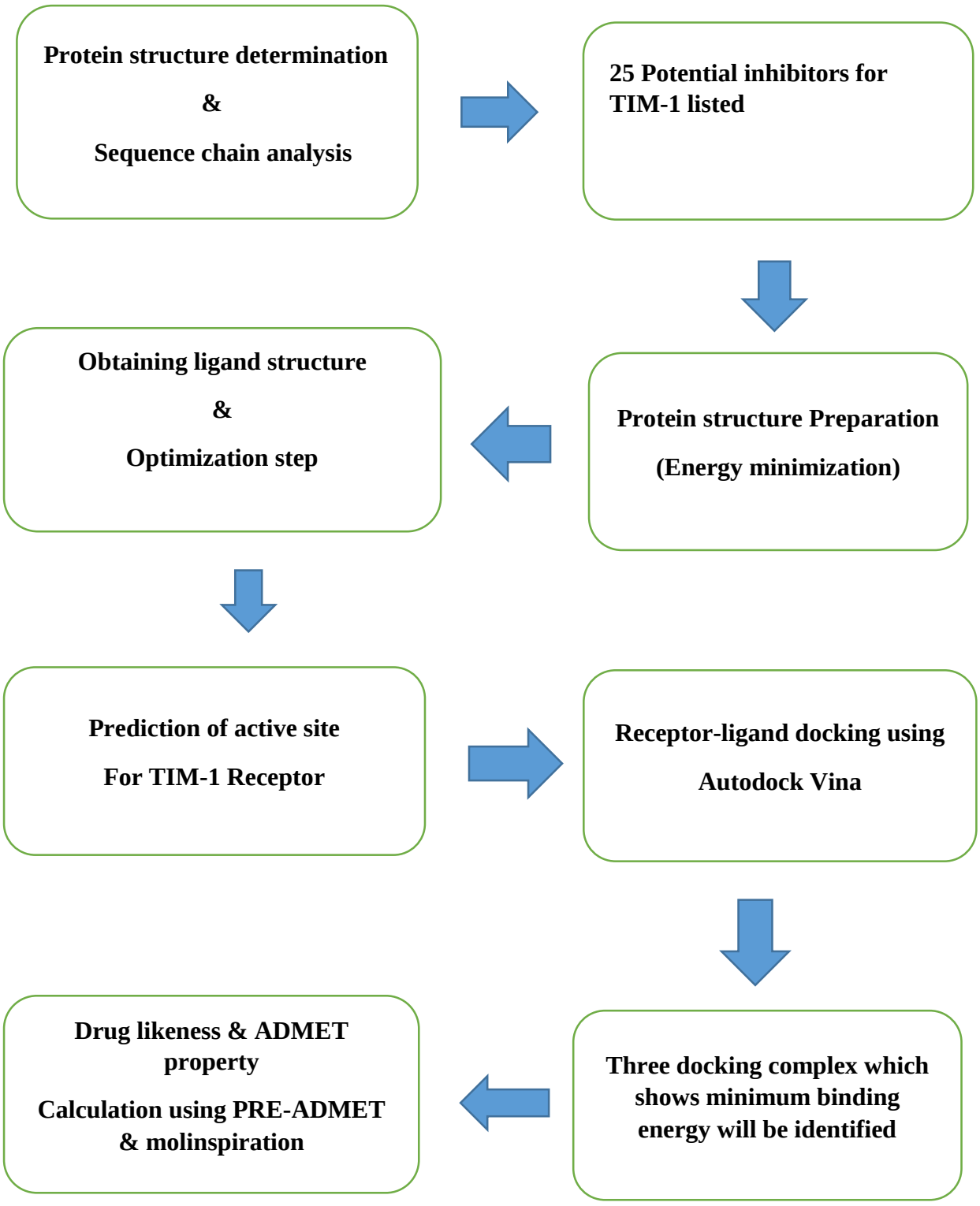
3.9. DISCOVERY STUDIO 4.1:

This is a visualization tool and have a vast application in molecular docking process for drug designing purpose. This software is very helpful for finding interacting residues during docking of ligand compounds with protein. [14]

CHAPTER 4

METHODOLOGY

4.METHODOLOGY



4.1. Protein structure determination and sequence chain analysis:

PDB ID of TIM-1 receptor is found to be 2OR8. Then the pdb website was accessed (www.pdb.org). Typed the PDB ID 2OR8 in search box, this got the result of sequence chain and checked the 3D view of the protein from PDB database on Jmol viewer. Then the Pdb file 2OR8 is downloaded and saved as .pdb format in the system.

. TIM-1 has a total of 2 chains. These chains are represented by unique entities of two sequences. Information mentioned below regarding the protein TIM-1 are determined from PDB.

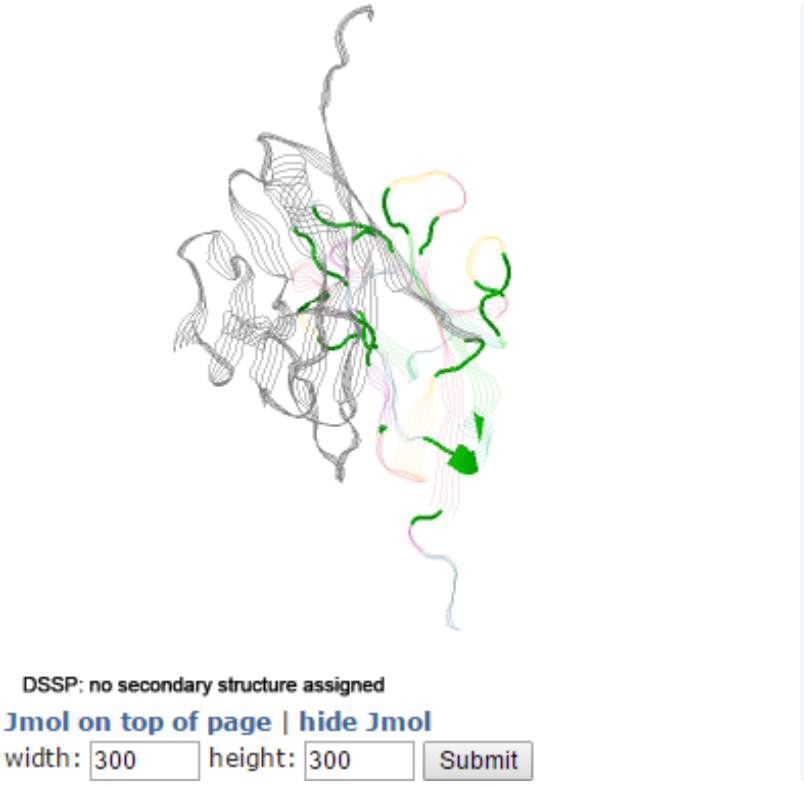


Figure 4.1 3D view of tim-1, source:PDB, image is viewed under Jmol [10]

CHAIN A:

This is the first chain (A). This light polymeric chain contains 47 residues. The secondary structure shown to be 5% helical with 2 helices and having 6 residues that contains 40% beta sheets and 11 strands. The secondary structure of this receptor protein is mentioned below:

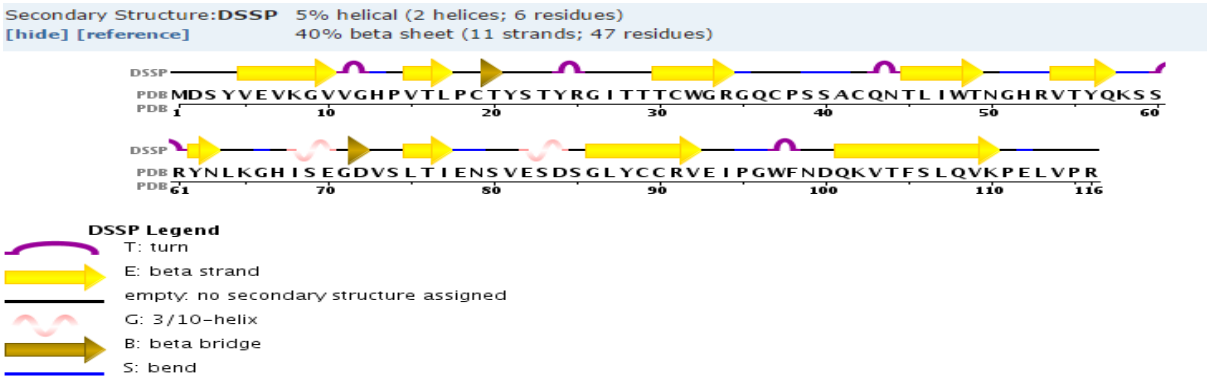


Figure 4.2 TIM-1 secondary structure, CHAIN A , Courtesy:PDB

CHAIN B:

This is the second chain (B). This polymeric chain contains 49 residues. The secondary structure shown to be 5% helical with 2 helices and having 6 residues that contains 42% beta sheets and 11 strands. The secondary structure of this receptor protein is mentioned below:

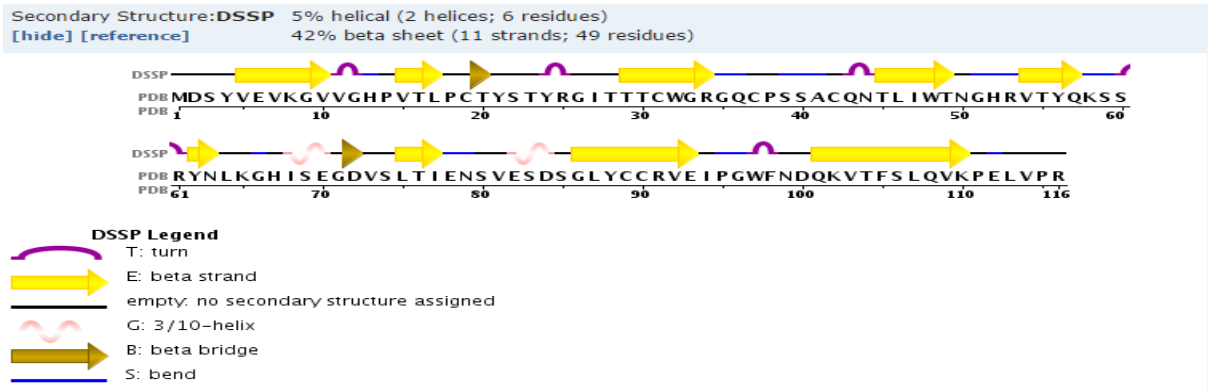
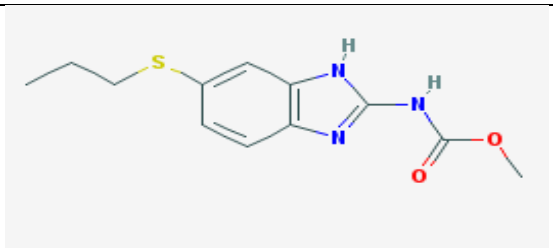
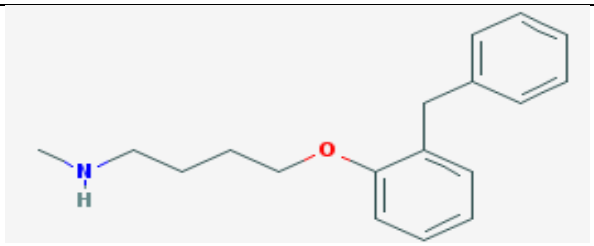
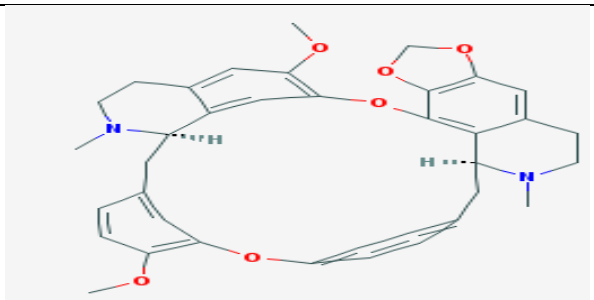


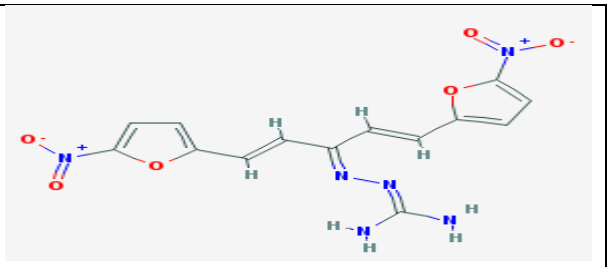
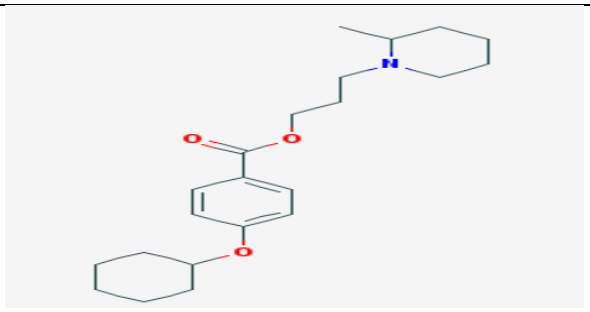
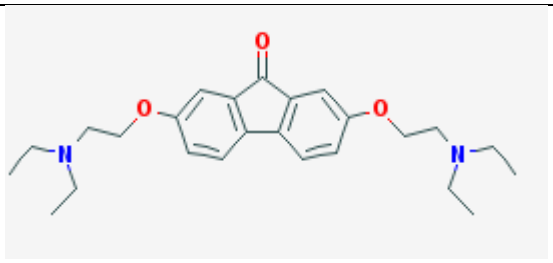
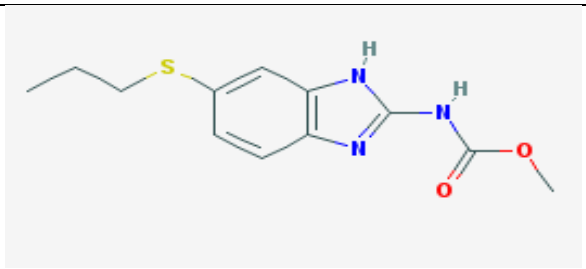
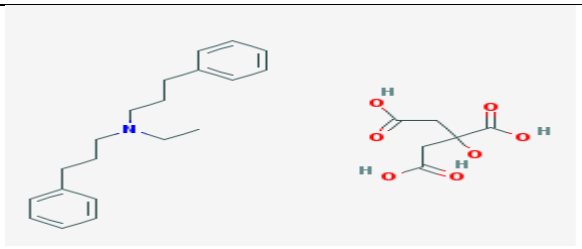
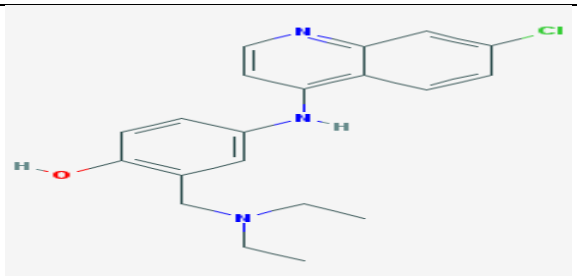
Figure 4.3 TIM-1 secondary structure, CHAIN B, Courtesy:PDB

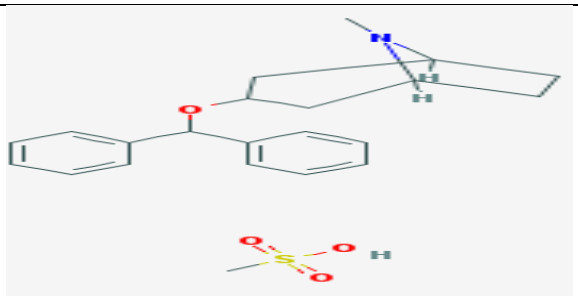
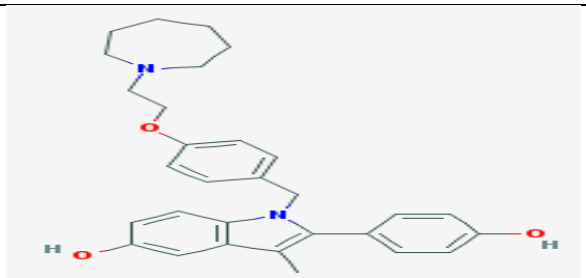
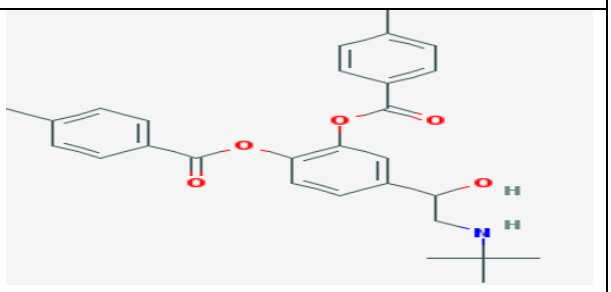
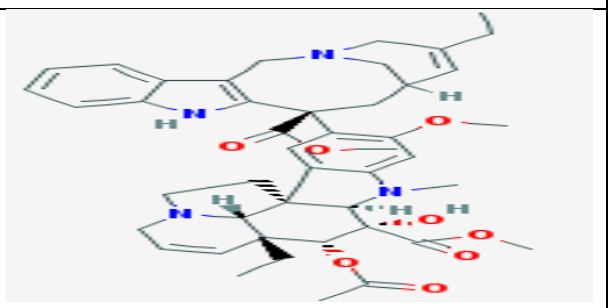
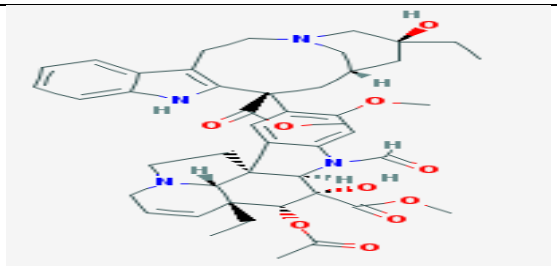
4.2. Inhibitor identification and preparation of their structure:

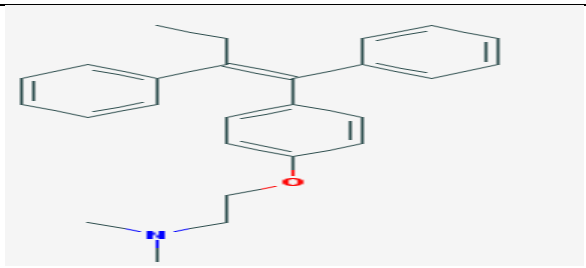
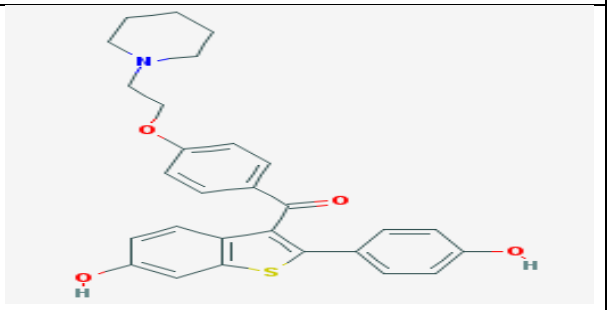
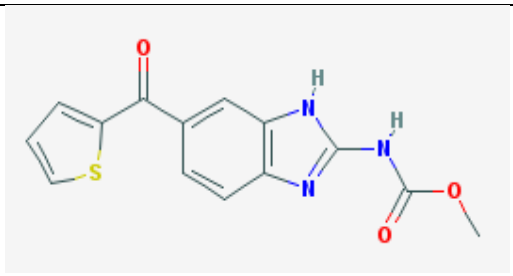
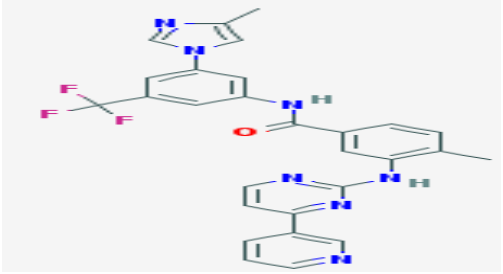
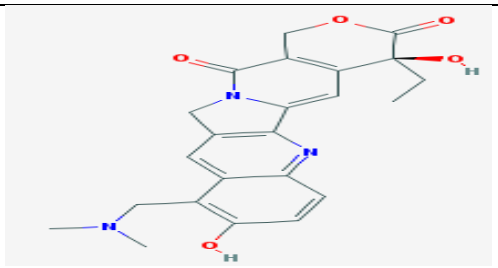
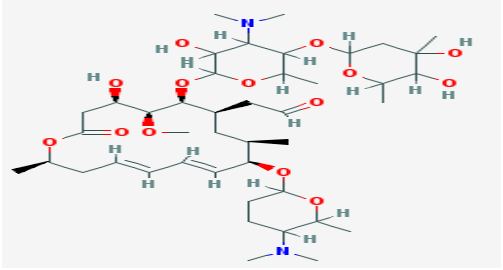
25 non-specific EBOV inhibitors were found out and listed. From the Nature Paper Journal, selection of 25 drug compounds was done. Downloading of corresponding chemical compound structure was done from PUBCHEM database. The files downloaded was in .sdf form. So Openbabel tool was used to change the format to .pdb file. A list of 25 selected inhibitor are listed below, those are to be docked with TIM-1 protein.

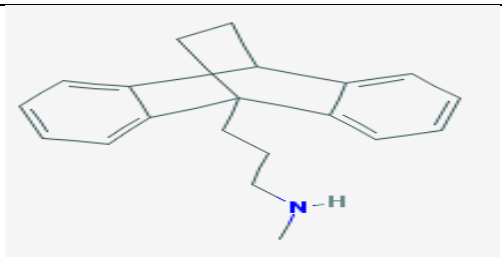
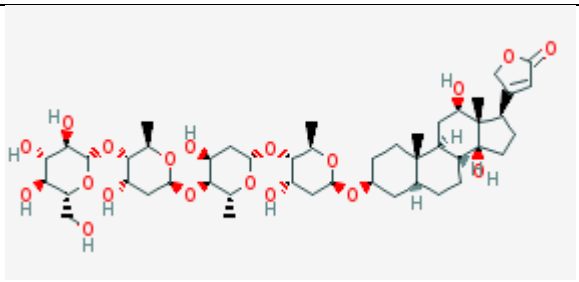
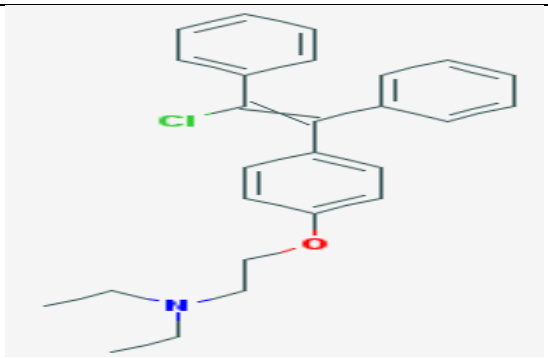
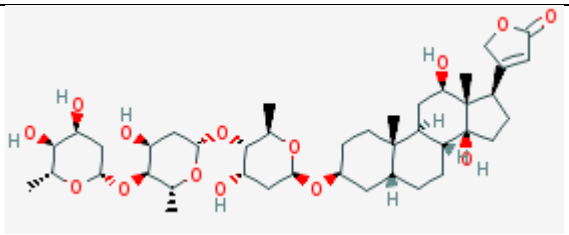
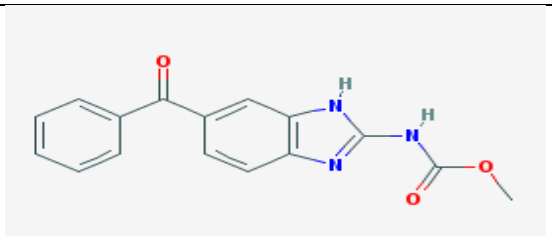
TABLE 4.1 List of inhibitors

1) AZACLORZINE Pubchem CID- 39526 Molecular Weight – 413.96 g/mol Molecular Formula- $C_{22}H_{24}Cl_1N_3O_5$	
2) BIFEMELANE Pubchem CID- 2377 Molecular Weight – 269.58 g/mol Molecular Formula- $C_{18}H_{23}NO$	
3) CEPHARANTHINE Pubchem CID- 10206 Molecular Weight – 606.707 g/mol Molecular Formula- $C_{37}H_{38}N_2O_6$	

4) NITROVIN Pubchem CID- 6252910 Molecular Weight – 360.28 g/mol Molecular Formula- $C_{19}H_{12}N_6O_6$	
5) CYCLOMETHYCAINE Pubchem CID- 10839 Molecular Weight – 359.50 g/mol Molecular Formula- $C_{22}H_{23}NO_3$	
6) TILORONE Pubchem CID- 5475 Molecular Weight – 410.54 g/mol Molecular Formula- $C_{25}H_{34}N_2O_3$	
7) ALBENDAZOLE Pubchem CID- 2082 Molecular Weight – 265.33 g/mol Molecular Formula- $C_{12}H_{15}N_3O_2S$	
8) ALVERINE CITRATE Pubchem CID- 21718 Molecular Weight – 473.55 g/mol Molecular Formula- $C_{26}H_{35}NO_7$	
9) AMODIAQUINE Pubchem CID- 2165 Molecular Weight – 355.86 g/mol Molecular Formula- $C_{20}H_{22}Cl_1N_3O$	

10) BENZOTROPINE MESYLATE Pubchem CID- 238053 Molecular Weight – 463.54 g/mol Molecular Formula- $C_{22}H_{29}NO_4S$	
11) BAZEDOXIFENE Pubchem CID- 154257 Molecular Weight – 470.60g/mol Molecular Formula- $C_{30}H_{34}N_2O_3$	
12) BITOLTEROL Pubchem CID- 35330 Molecular Weight – 461.54g/mol Molecular Formula- $C_{28}H_{31}NO_5$	
13) VINORELBINE Pubchem CID- 60780 Molecular Weight – 778.93g/mol Molecular Formula- $C_{45}H_{54}N_4O_8$	
14) VINCRISTINE Pubchem CID- 5978 Molecular Weight – 824.95g/mol Molecular Formula- $C_{46}H_{56}N_4O_{10}$	

15) TAMOXIFENE Pubchem CID- 2733526 Molecular Weight – 371.54g/mol Molecular Formula- $C_{26}H_{29}NO$	
16) RALOXIFENE Pubchem CID- 5035 Molecular Weight – 473.58g/mol Molecular Formula- $C_{28}H_{27}NO_4$	
17) NOCODAZOLE Pubchem CID- 4122 Molecular Weight – 301.32g/mol Molecular Formula- $C_{14}H_{11}N_3O_3S$	
18) NILOTINIB Pubchem CID- 644241 Molecular Weight – 529.515g/mol Molecular Formula- $C_{28}H_{22}N_7OF_3$	
19) TOPOTECAN Pubchem CID- 60700 Molecular Weight – 421.44g/mol Molecular Formula- $C_{23}H_{23}N_3O_5$	
20) SPIRAMYCIN Pubchem CID- 6419898 Molecular Weight – 843.05g/mol Molecular Formula- $C_{43}H_{74}N_2O_4$	

21) MAPROTILINE Pubchem CID- 4011 Molecular Weight – 277.40g/mol Molecular Formula- $C_{20}H_{23}N$	
22) DESLANOSIDE Pubchem CID- 28620 Molecular Weight – 943.07g/mol Molecular Formula- $C_{47}H_{74}O_{19}$	
23) CLOMIPHENE Pubchem CID- 2800 Molecular Weight – 405.95 g/mol Molecular Formula- $C_{26}H_{28}ClNO$	
24) DIGOXIN Pubchem CID- 2724385 Molecular Weight – 780.93g/mol Molecular Formula- $C_{41}H_{64}O_{14}$	
25) MEBENDAZOLE Pubchem CID- 4030 Molecular Weight – 295.29g/mol Molecular Formula- $C_{16}H_{13}N_3O_4$	

4.3. Prediction of Active site :

Before docking the active site or the binding site of the protein was predicted. There are so much online tools and softwares are available to do the prediction of active site. CastP is one of the best software to evaluate the binding region or site in a protein.

First of all, for active site calculation CastP requires the pdb file of the protein. As soon as the file has been uploaded or the PDB code is searched, it will give the all details about the active site. Then the result was viewed using several tools for purpose of visualization such as PYMOL, discovery studio etc. It gives the Cartesian X, Y, Z coordinates i.e. active site. These values will be used futher while docking experiments will start to happen in Autodock Vina. Grid Box prepation usually done prior to the docking step, so at that time it requires these coordinate values.

4.4. Protein and ligand optimization :

Firstly the structure of the TIM-1 protein was acquired from the PDB site. Using the PDB ID (2OR8), the structure was downloaded and viewed in the Visualization tool PYMOL. Before performing the docking, some steps has been done known as minimization of energy or optimization step. In Vina, most PDB structures do not have hydrogen. So we need to add them using ADT. Structure has been minimized by adding polar hydrogen atom and removing water molecules. Further charges have been added such as kollman and gasteiger charge as for filling the missing loops. Now a charge field has been set up by adding polar hydrogen atoms and removing unwanted water. Again the ligands also downloaded from the PUBCHEM website and Drugbank library. The compounds downloaded are in sdf format and that has to be changed to pdb as this is prerequisite for ADT.

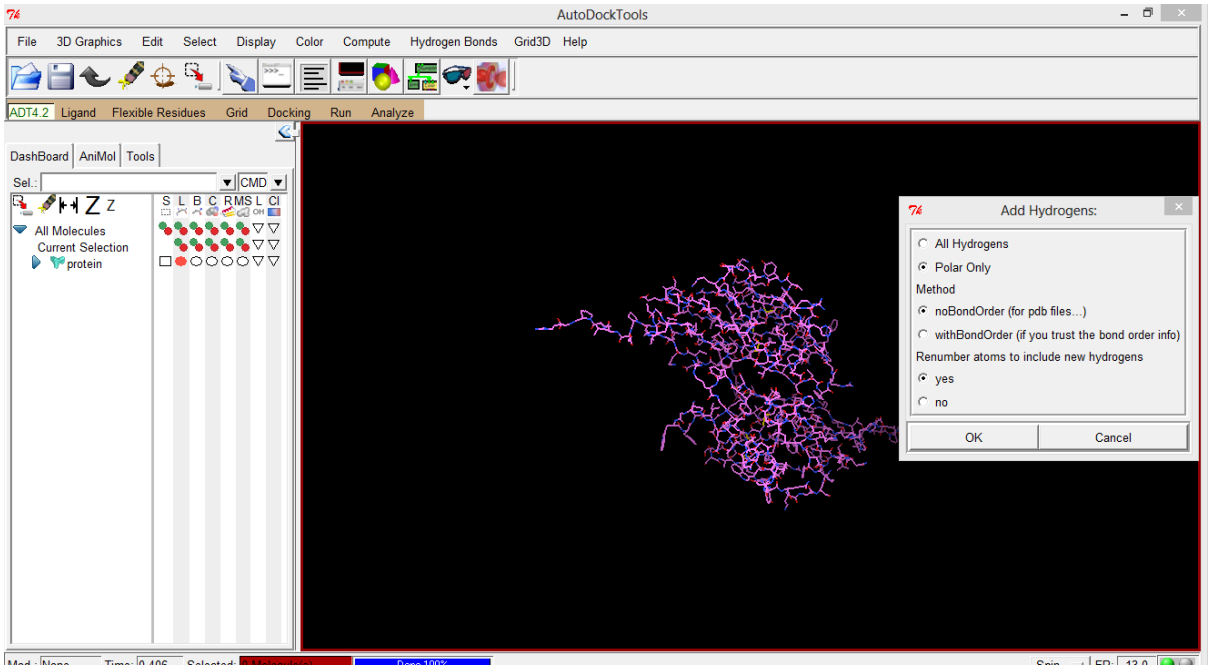


Figure 4.4 screenshot during addition of hydrogen atom

Total kollman charges added = 8.294

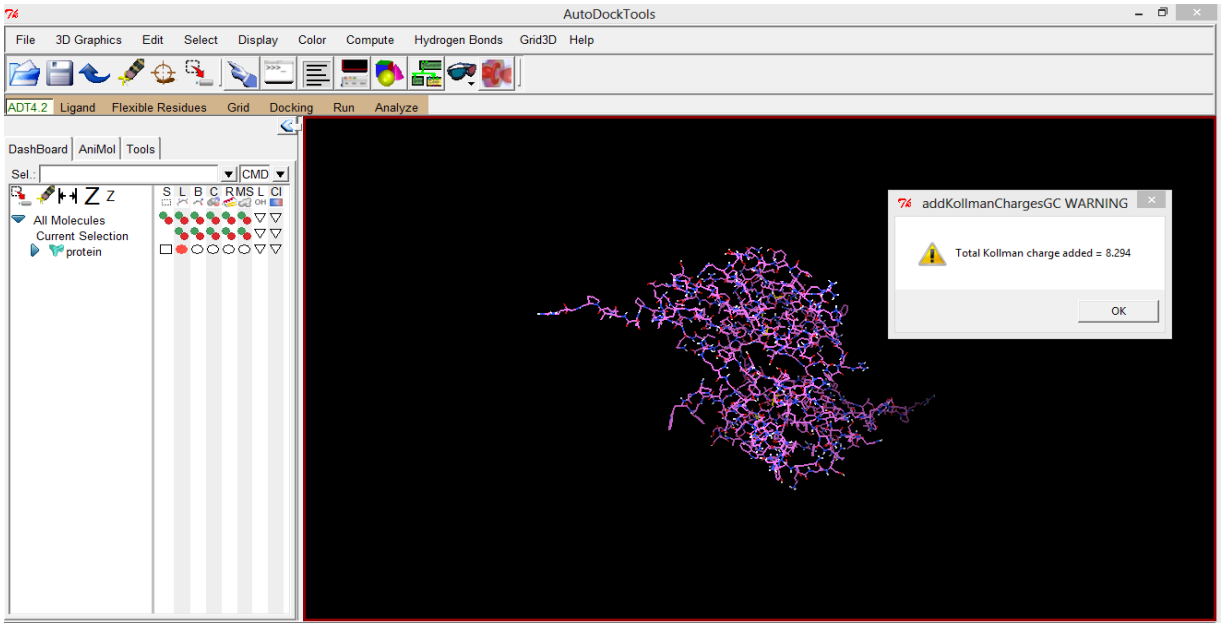


Figure 4.5 screenshot while adding the kollman charge

Total gasteiger charge added = 0.998

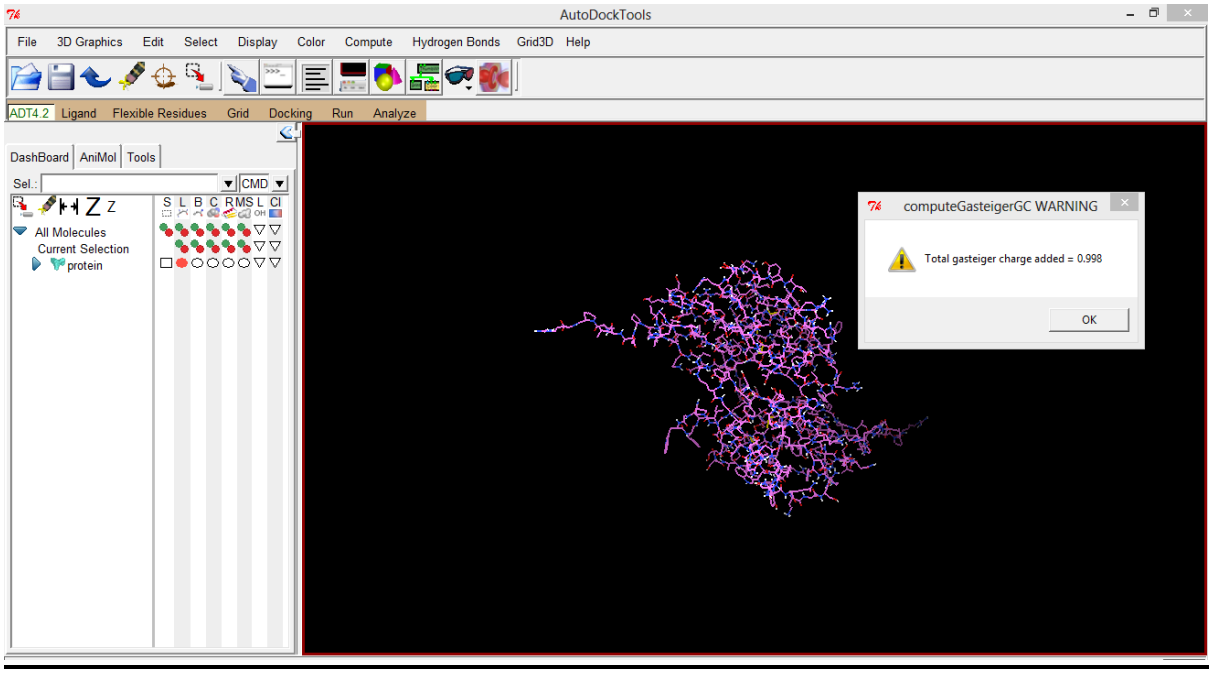


Figure 4.6 screenshot while adding the gesteiger charge

4.5. Preparation of Grid Box:

A Grid Box has to be set up around the binding site or active site of the protein. It is a necessary step to perform molecular docking of ligand and receptor. By clicking on the grid option on the ADT interface, it will show a box for grid option. The coordinate values and the default parameters are entered. Number of points have been modified in X, Y, Z dimension by rotating the dial. The grid box should be large enough to accommodate the ligand in its extended conformation. As far as possibly it should not go into the empty space.

In Grid option box:

Number of points in X-dimension = 94

Number of points in Y-dimension = 40

Number of points in Z-dimension = 50

Centre Grid Box:

X centre = -7.385

Y centre = 17.199

Z centre = -18.915

After setting this parameter values spin option has been choosed and turned it ON. It will start spinning on the screen showing all parts of protein from all different angles whether it has covered the site properly or not.

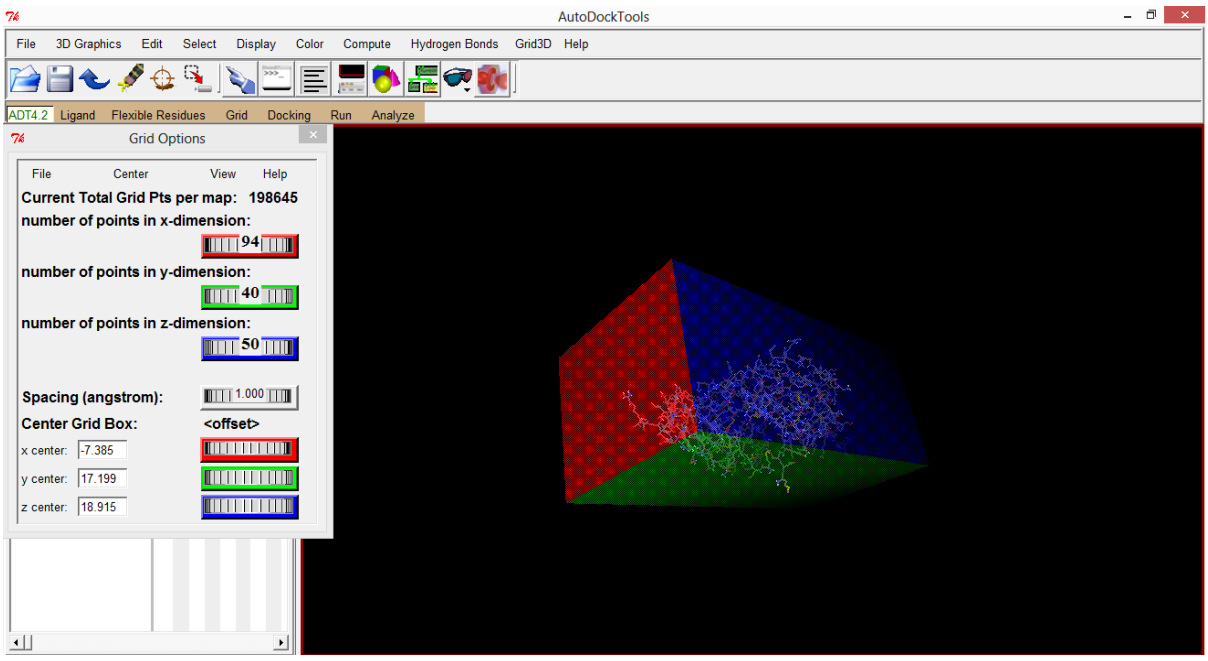


Figure 4.7 screen shot during Grid Box Preparation

4.6. Docking using VINA :

Docking is a method which evaluates the position of inhibitor compounds with the receptor for prediction of binding energy of complex or the affinity of ligands towards the protein. Docking analysis has been done with the help of Autodock Vina. Before that the optimization of protein and ligand was done. This two files were in .pdb format and placed in a folder named by the respective ligand. 25 compounds have been used, so 25 folders have been created with giving the ligand names. Then the protein molecule was saved in .pdbqt format and so was for the ligand also. A conf.txt file also included in the folder. It has a certain format to give the input text. Now it was all set for docking.

conf.txt File :

receptor = protein.pdbqt
ligand = ligand.pdbqt

center_x = -7.385
center_y = 17.199
center_z = 18.915

size_x = 94
size_y = 40
size_z = 50
out = result.pdbqt

Command Prompt:

For running the docking simulation and getting the binding affinity values certain commands were given. As the commands put properly in the command prompt affinity values will be coming for the respective ligand. 9 modes were evaluated and the first mode was taken as the best conformation of the compound in the binding site.

This same procedure was repeated for other compounds left for docking purpose. So a series of docking experiments was performed and value of least binding energy mode was noted down. Then log.txt file contains all the affinity scores of 9 conformations of the ligand. Result.pdbqt file was viewed using visualization tools like discovery studio and PYMOL.

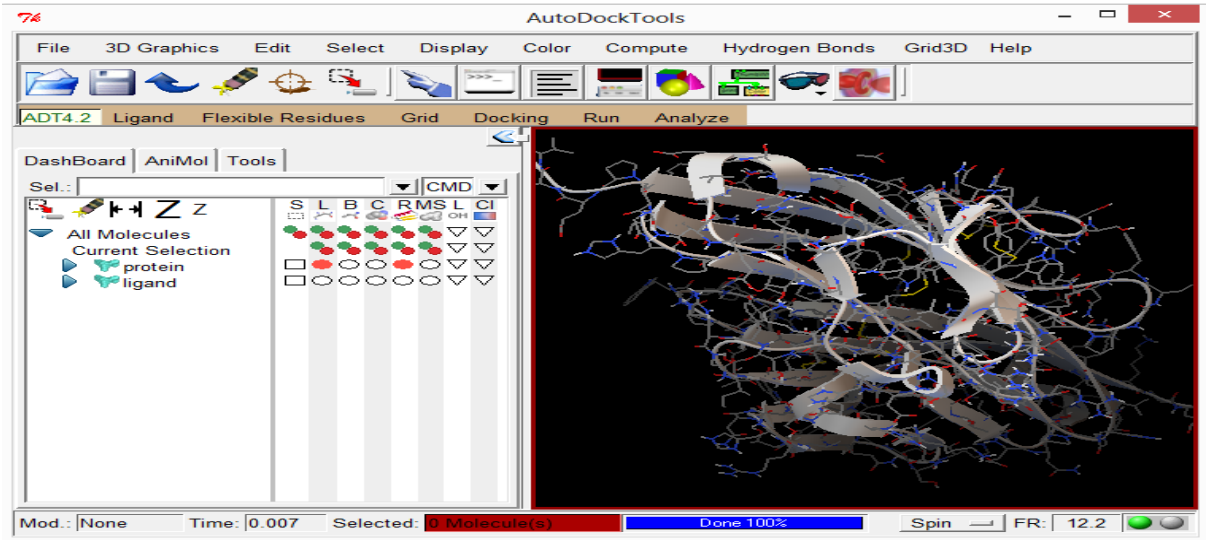


Figure 4.8 screenshot during performing docking with ADT

4.7. Bioactivity & ADMET properties Evaluation:

Docking was performed and three binding complex having minimum energy will be essentially separated. The Bioactivity, physiochemical properties were checked from the website of molinspiration[11]. It gave all the results with showing properties of respective lead compounds. ADMET properties calculated from Pre-ADMET server. From this the best ligand compound was chose.

CHAPTER 5

RESULTS & DISCUSSION

5.RESULTS & DISCUSSION

5.1. Minimization of Protein :

Before Docking the PDB structure of protein TIM-1 was downloaded from the website (PDB ID- 2OR8). Minimization of Protein was done by adding polar H-atoms and removing water molecules which are not wanted. The process follows by addition of Hydrogen and removal of water parts for better binding.

Total kollman charges added = 8.294

Total kollman charges added = 0.998

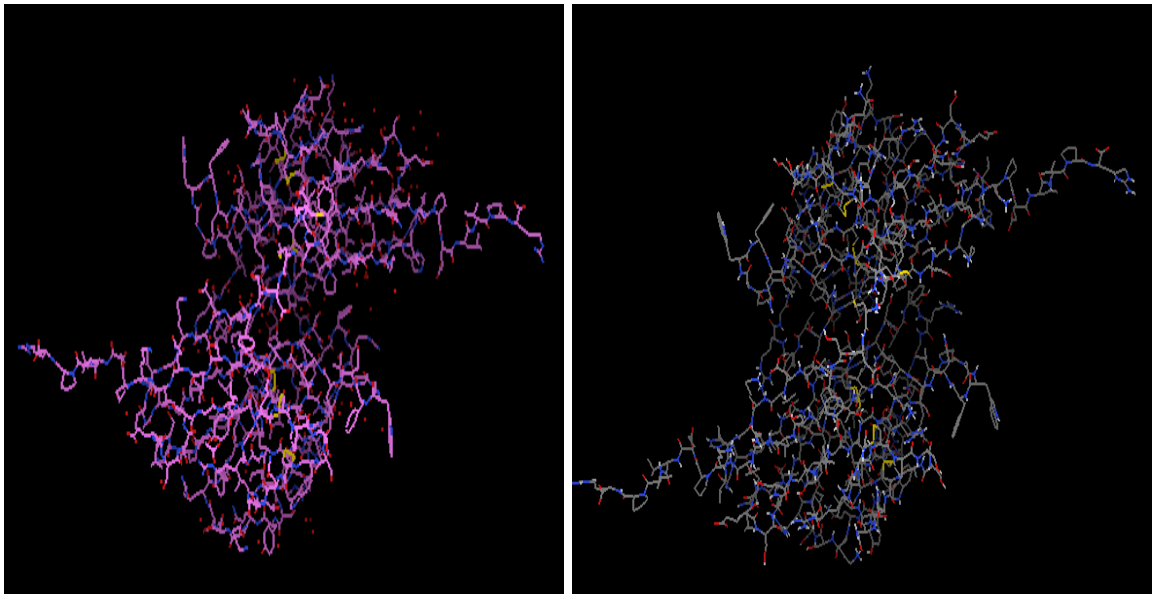
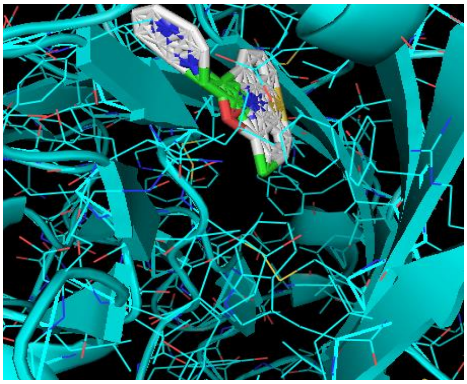


Figure 5.1 Protein before minimization and after minimization

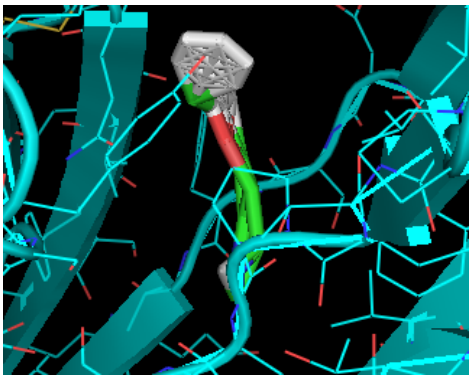
5.2. Docking Results:

After Docking of 25 compounds, 9 orientaion values or modes were generated for each ligands. The first mode or conformation was considered as the best binding residue. The value of the first mode was noted down for that ligand. Thus a series of docking was performed so as to get the corresponding binding affinity scores. Binding complex of ligand with receptor was shown in PYMOL and DS 4.1.

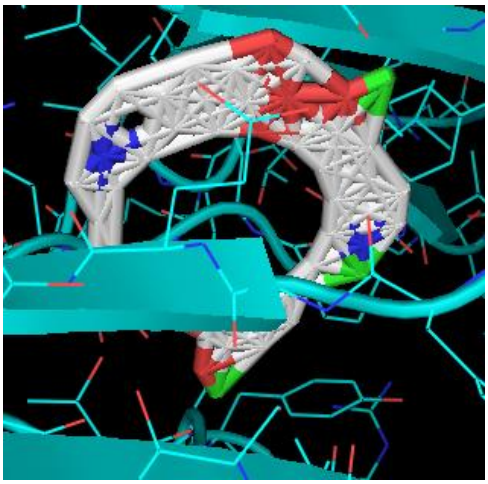
5.2.1. Docked complex structures:



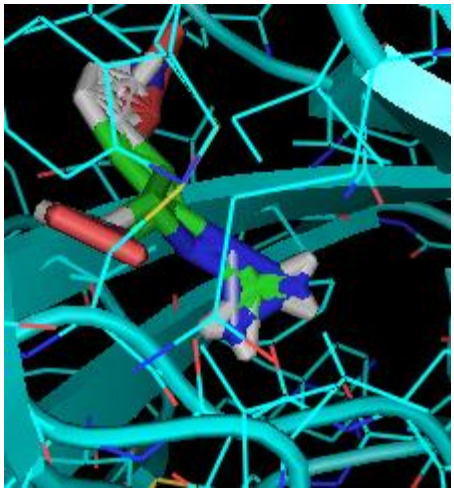
(1)



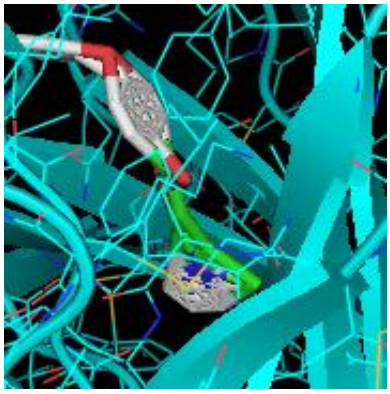
(2)



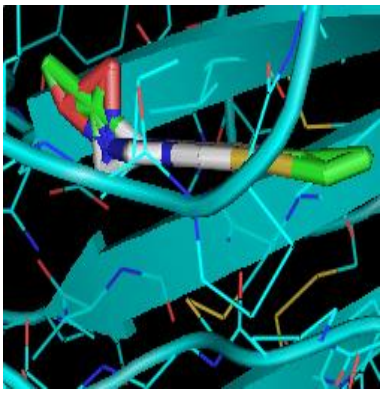
(3)



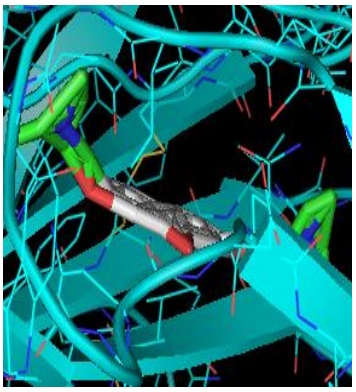
(4)



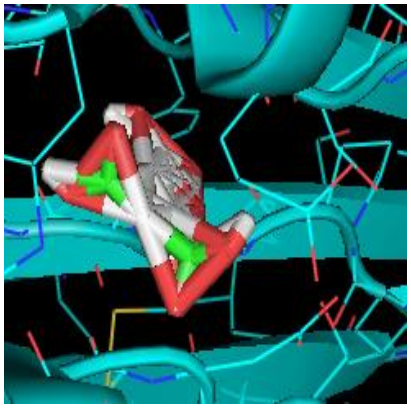
(5)



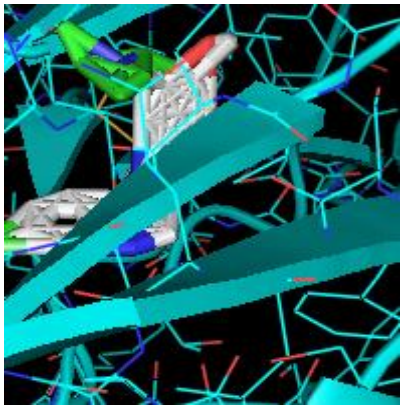
(6)



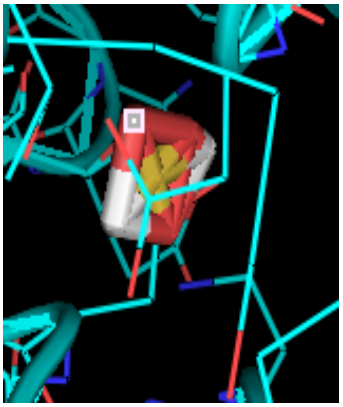
(7)



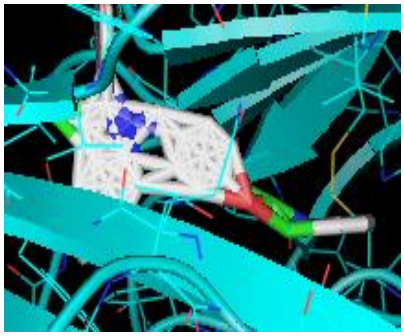
(10)



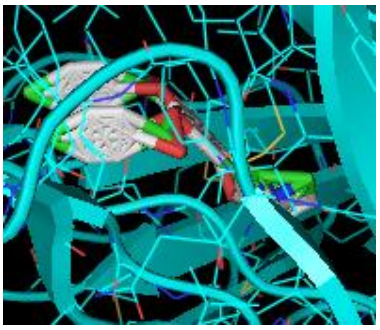
(11)



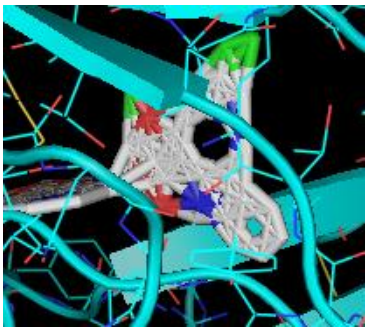
(12)



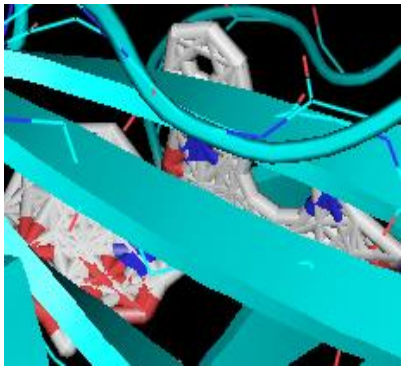
(13)



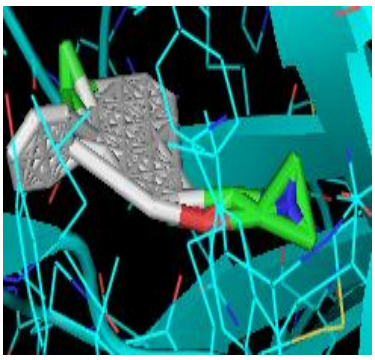
(14)



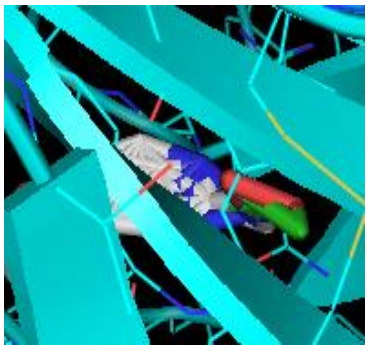
(15)



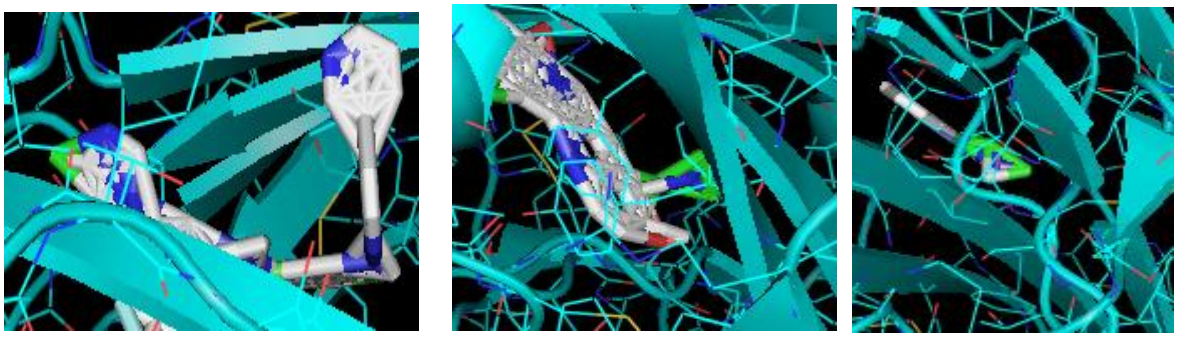
(16)



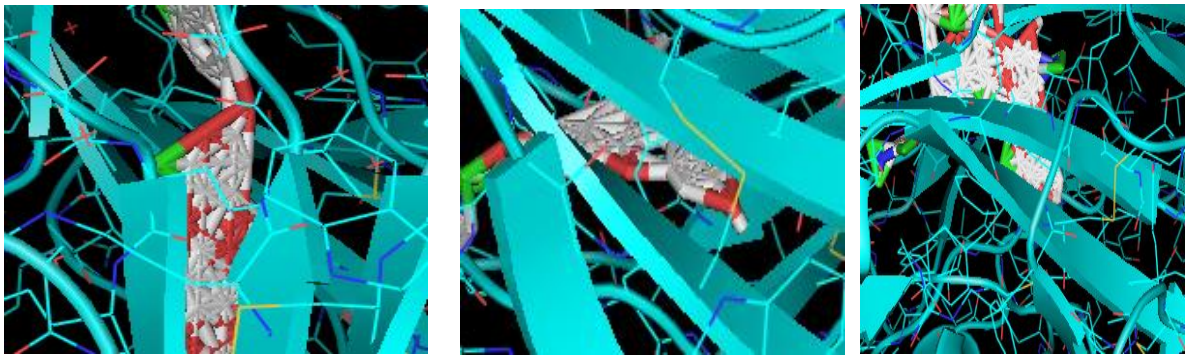
(17)



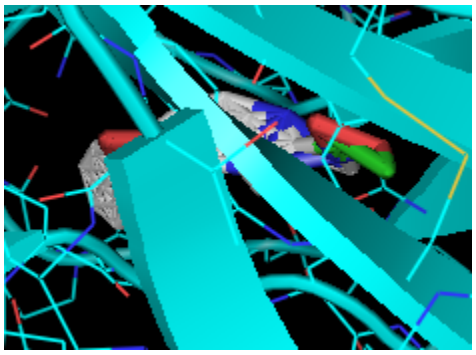
(18)



(19) (20) (21)



(22) (23) (24)



(25)

Figure 5.2 Docked complex of all ligands (1-25)

5. 2.2. **Binding Energy Values of the ligands:**

Table 5.1 Affinity value of all ligands

Ligands	**Affinity value (kcal/mol)**
Azaclorzine	-9.8
Bifemelane	-6.6
Cepharanthine	-13.3
Nitrovin	-7.3

Tilorone	-7.4
Albendazole	-6.5
Alverine Citrate	-8.3
Amodiaquine	-7.3
Benzotropine mesylate	-4.2
Bazedoxifene	-10.7
Bitolterol	-8.7
Vinorelbine	-15.9
Vincristine	-14.3
Tamoxifene	-8.7
Raloxifene	-11.2
Nacodazole	-8.0
Nilotinib	-11.9
Topotecan	-10.1
Spiramycine	-12.8
Maprotiline	-7.9
Deslanoside	-16.7
clomifene	-9.0
Digoxin	-15.7
Mebendazole	-8.0

From the above table, it is shown that Deslanoside, Vinorelbine and Digoxin having the least binding energy value and forming the most stable docking complex with the protein in the active site. These compound have the better binding affinity compared to others. Among these three, Deslanoside has the highest dock score of -16.7 kcal/mol so this is having the best interaction among all the compounds docked with the protein. Likewise Vinorelbine acquires a dock score of -15.9 kcal/mol which is better affinity value in comparison to Digoxin. The binding complex of these ligands with the protein visualized in Disocvery Stusio version 4.1. It shows all 9 interacting residues of the ligands within the binding site. The properties of these ligands also checked using molinspiration and Pre-ADMET. Below a binding energy plot has delineated.

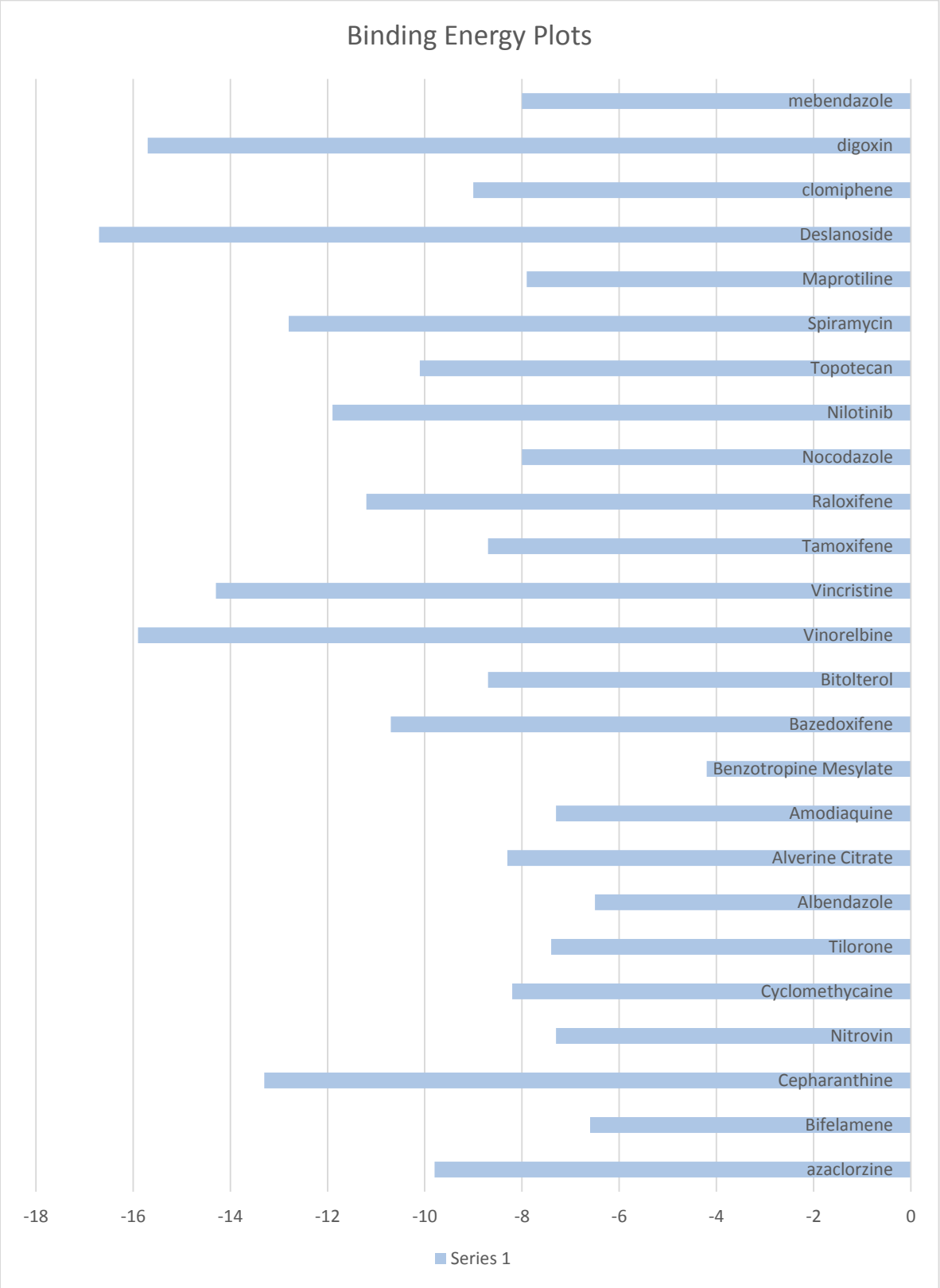
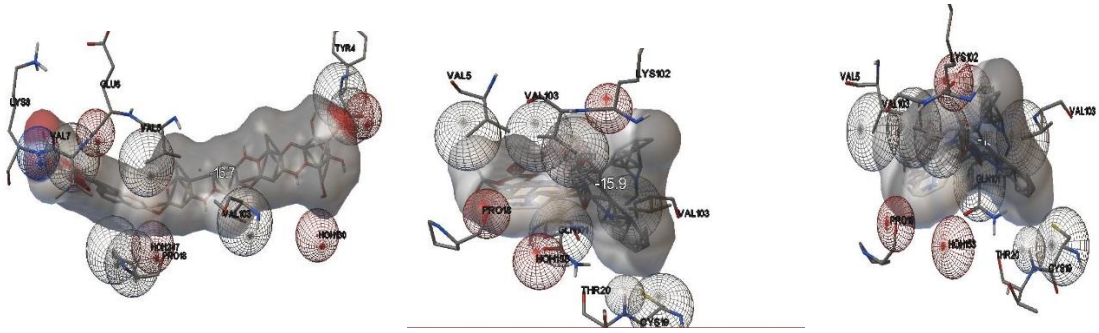


Figure 5.3 Binding Energy plot of all ligands



(Deslanoside) **(vinorelbine)** **(digoxin)**

Figure 5.4 binding complex with residues using Discovery Studio

Table 5.2 ligands with interacting residue name

Ligand	Affinity (kcal/mol)	Interacting Residues
DESLANOSIDE	**-16.7**	**LYS8,VAL7,GLU6,VAL5,HOH247,PRO18,VAL103, TYR4,HOH130**
VENORELBINE	**-15.9**	**VAL5,VAL103,PRO18,HOH158,GLN101,THR20,CYS19,VAL103,LYS102**
DIGOXIN	**-15.7**	**PRO18,VAL5,HOH158,VAL103,LYS102,VAL103,LYS102,GLN101, THR20**

5.3. Bioactivity and physiochemical properties:

Table 5.3 Bioactivity and physiochemical Properties

Properties	Deslanoside	Vinorelbine	Digoxin
LogP	-0.808	5.918	1.117
TPSA	282.229	133.882	203.077
Rotatable Bonds	10	10	7
GPCR ligand	-3.26	-1.31	-1.28
Kinase inhibitor	-3.91	-2.55	-2.17
Ion ch. modulator	3.82	-2.67	-2.39
Protease inhibitor	-2.89	-1.20	-0.88
Nuclear receptor	-3.71	-2.39	-1.88
Enzyme inhibitor	-3.09	-2.00	-1.26
nON	19	12	14
nOHNH	9	2	6
IC_{50}(μM)	0.485	0.066	0.763
IC_{90} (μM)	11.7	0.190	3.45
Max Inh.(%)	66	90	68
CytotoxicityIC_{50}(μM)	250	250	>500
SelectivityIndex(fold)	515	>7546	327
MOA	Na^+-K^+ pump inhibitor	Microtubule Inhibitor	Na^+-K^+ pump Inhibitor

ADMET Prediction:

Deslanoside:

ADME PreADMET

	ID	Value
	BBB	0.0313231*
	Caco2	19.0856
	CYP_2C19_inhibition	Non
	CYP_2C9_inhibition	Inhibitor
	CYP_2D6_inhibition	Non
	CYP_2D6_substrate	Non
	CYP_3A4_inhibition	Inhibitor
	CYP_3A4_substrate	Substrate
	HIA	8.852000
	MDCK	0.0230705*
	Pgp_inhibition	Non
	Plasma_Protein_Binding	39.011827
	Skin_Permeability	-4.89328*
	ID	Value

(a)

Digoxin:

ADME PreADMET

	ID	Value
	BBB	0.0597861
	Caco2	19.9372
	CYP_2C19_inhibition	Non
	CYP_2C9_inhibition	Inhibitor
	CYP_2D6_inhibition	Non
	CYP_2D6_substrate	Non
	CYP_3A4_inhibition	Inhibitor
	CYP_3A4_substrate	Substrate
	HIA	58.132712
	MDCK	0.0448262
	Pgp_inhibition	Inhibitor
	Plasma_Protein_Binding	63.806209
	Skin_Permeability	-4.98301
	ID	Value

(b)

Venorelbine:

ADME PreADMET

	ID	Value
	BBB	0.73315
	Caco2	40.6048
	CYP_2C19_inhibition	Non
	CYP_2C9_inhibition	Inhibitor
	CYP_2D6_inhibition	Non
	CYP_2D6_substrate	Weakly
	CYP_3A4_inhibition	Inhibitor
	CYP_3A4_substrate	Substrate
	HIA	95.946339
	MDCK	0.0434155*
	Pgp_inhibition	Inhibitor
	Plasma_Protein_Binding	64.298578
	Skin_Permeability	-3.66733*
	ID	Value

(c)

Figure 5.5 ADMET properties of (a) Deslanoside (b) Digoxin (c) Venorelbine

CHAPTER 6

CONCLUSION

6.CONCLUSION

The current study focused on molecular docking analysis of 25 potential inhibitors of Ebola virus with the TIM-1 receptor. The main goal was to propose the most favourable ligand compound that could be effective to target the protein i.e. TIM-1. From the result it was determined that Deslanoside, Vinorelbine and Digoxin formed stable complex compared to other compounds. These compounds showed maximum binding affinity to the receptor. These could be the best compounds to inhibit the Ebola entry into the host cell. Deslanoside and digoxin have a function to inhibit Na^{+}-K^{+} pump as well as from the determined ADMET properties these can be accepted as potential therapeutic for Ebola infection. More in vivo study is required to analyse the mechanism of this virus and interaction mechanism of the compounds which are shown in this study. So the main aim is to narrow down the efforts of researchers in clinical laboratories by proposing these compounds for wet lab study which may lead to identify the novel anti-viral agent to inhibit Ebola infection.

REFERENCES

REFERENCES:

1) Heinz F, Geisbert TW: Ebola hemorrhagic fever. *The Lancet* 2011, 377:849-862
2) Casillas AM1, Nyamathi AM, Sosa A, Wilder CL, Sands H. A current review of Ebola virus: pathogenesis, clinical presentation, and diagnostic assessment. Biol Res Nurs.2003; 4(4): 268-75.
3) Timmins J, Schoehn G, Richard BS: Ebola virus matrix protein VP40 interaction with human cellular factors Tsg 101 and Nedd 4. *J Mol Biol* 2003, 326: 493-502
4) Ebola and Marburg viruses: molecular and cellular biology. *(Horizon Bioscience, 2004)*
5) Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J. & Nichol, S. T. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 3602–3607 (1996).
6) J. E. Lee, M. L. Fusco, A. J. Hessell, W. B. Oswald D. R. Burton & E. O Saphire (2008) Structure of the ebola virus glycoprotein bound to an antibody from ahuman survivor. *Nature* **454**, 177-182.
7) Ascenz P: Ebola virus and Marburg virus: Insight the Filoviridae Family. *Mol Aspects Med* **2008**, 29: 151-85
8) Dolnik O, Kolesnikova L, Becker S. 2008. Filoviruses: interactions with the host cell. Cell. Mol. Life Sci. 65:756–776]
9) Rennert PD. 2011. Novel roles for TIM-1 in immunity and infection.Immunol. Lett. 141:28–35.

10) Michael F. sannerv python : A programming language for software integration and development J. Mol Graphics Mod , 1999, vol-17, pp 57-61

11) Sven Moller-Tank, Andrew S. Kondratowicz, Robert A. Davey, Paul D. Rennert and Wendy Maury, 2013, Role of the Phosphatidylserine Receptor TIM-1 in Enveloped-Virus Entry

12) Molinspiration:www.molinspiration.com.

13) Joe Dundas, Zheng Ouyang, Jeffery Tseng, Andrew Binkowski, Yaron Turpaz, JieLiang.2006. CASTp: computed atlas of surface topography of proteins with structural andtopographical mapping of functionally annotated residues Nucleic Acids Research, Vol.34, No. suppl 2. (1 July), pp. W116-W118, doi:10.1093/nar/gkl282.

14) Accelrys Software Inc.2013. Discovery Studio Modeling Environment, Release 4.0, SanDiego: Accelrys Software Inc.

15) Hoenen T, Groseth A, Kolesnikova L, Theriault S, Ebihara H, Hartlieb B, Bamberg S, Feldmann H, Stroher U, Becker S: Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. *J Virol* 2006, 80:7260–7264

16) Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR, Saphire EO: Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 2008, 454:177-182

17) Baron RC, McCormick JB, Zubeir OA. Ebola hemorrhagic fever in southern Sudan: hospital dissemination and intra familial spread. Bull WHO 1983; 6: 997-1003.

18) F. X. Gomis-Ruth, A. Dessen, J. Timmins, A.Bracher, L. Kolesnikowa, S. Becker, H. D. Klenk & W.Weissenhorn (2003)

19) Hartlieb, B. & Weissenhorn, W. Filovirus assembly and budding. *Virology* **344,** 64–70 (2006).